

**Snow hydrology signatures for model identification within a limits-of-acceptability approach**

Journal:	<i>Hydrological Processes</i>
Manuscript ID	HYP-15-0871.R1
Wiley - Manuscript type:	Keith Beven Tribute
Date Submitted by the Author:	n/a
Complete List of Authors:	Schaefli, Bettina; Ecole Polytechnique Fédérale de Lausanne, Laboratory of ecohydrology ECHO; Université de Lausanne, Institute of Earth Surface Dynamics
Keywords:	rainfall-runoff model, snow hydrology, model calibration, limits-of-acceptability, Monte Carlo simulation, Discharge signature

SCHOLARONE™  
Manuscripts

Version accepted for publication

**Snow hydrology signatures for model identification within a limits-of-acceptability approach**

(running title: Snow hydrology signatures)

Bettina Schaefli<sup>1,2</sup>

1: Laboratory of Ecohydrology and Laboratory of Hydraulic Constructions, School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL)

2: Now at: Faculty of Geosciences and Environment, University of Lausanne, 1015 Lausanne, Switzerland, tel: 0041 21 692 44 00

**Abstract**

Discharge simulation from snow-dominated catchments seems to be an easy task. Any spatially-explicit precipitation-runoff model coupled to a temperature-index snow model generally yields simulations that mimic well the observed daily discharges. The robustness of such models is, however, questionable: in presence of strong annual discharge cycles, small model residuals do not guarantee high explanatory power of the underlying model. This paper proposes a methodology for snow hydrological model identification within a limits-of-acceptability framework, where acceptable model simulations are the ones that reproduce a set of signatures within an a priori specified range. The signatures proposed here namely include the relationship between the air temperature regime and the discharge regime, a new snow hydrology signature that can be readily transferred to other Alpine settings. The discriminatory power of all analyzed signatures is assessed with a new measure of their discriminatory power in the model prediction domain. The value of the proposed snow hydrology signatures and of the limits-of-acceptability approach is demonstrated for the Dischma river in Switzerland, whose discharge shows a strong temporal variability of hydrologic forcing conditions over the last 30 years. The signature-based model identification for this case study leads to the surprising conclusion that the observed discharge data contains a multi-year period that cannot be reproduced with the model at hand. This model-data mismatch might well result from a yet to be identified problem with the discharge observations, which would have been difficult to detect in a classical residual-based model identification approach. Overall, the detailed results for this case study underline the robustness of the limits-of-acceptability approach in the presence of error-prone observations if it is applied in combination with relatively robust signatures. Future work will show whether snow hydrology signatures and their limits-of-acceptability can be regionalized to ungauged catchments, which would make this model selection approach particularly powerful for Alpine environments.

Key words (max. 6): rainfall-runoff model, snow hydrology, model calibration, limits-of-acceptability, Monte Carlo simulation, signature

## 1 Introduction

Streamflow modeling in high mountainous environments faces the fascinating challenge of dealing with water stored at different aggregation states within the soil, within the seasonal or the intermittent snow cover or within perennial snow and ice packs. The dimensionality gap (Beven and Binley, 1992; Gupta et al., 2008) between the number of processes occurring at different spatio-temporal scales to be described by a model and the number of time series available to evaluate or to calibrate the model might appear to be far more discouraging than for streamflow modeling in lowland systems.

The resulting problems of model equifinality discussed by Keith Beven since his 1993 paper (Beven, 1993) might indeed be exacerbated here, in particular because the streamflow signal observed for any catchment larger than a few square kilometres might efficiently average out any spatial heterogeneity of snow accumulation and of melt release (Comola et al., 2015b).

On the other hand, the temporal accumulation of frozen water on the surface offers also a great opportunity for parameter estimation since the presence of snow or ice is relatively easy to observe with methods ranging from point stake measurements to satellite data. An important amount of research addresses for example the question of how to use remote sensing data and in particular MODIS snow cover information to reduce the equifinality of snow hydrological models. Most of these studies compare distributed snow cover simulations to the presence and absence of snow in MODIS images (e.g. Finger et al., 2011; Parajka and Blöschl, 2008), but there are also more recent attempts to directly estimate snowmelt parameters from MODIS snow cover information (He et al., 2014) (Finger et al., 2011; He et al., 2014; Parajka and Blöschl, 2008).

Still largely unexplored is, however, the question of how to efficiently extract information on the dominant hydrological processes in high mountainous catchments from available discharge observations within a so-called diagnostic model identification framework (Yilmaz et al., 2008), where model identification can be understood in the words of Wagener and Gupta (2005) as “the process of identifying one or more suitable models for a specific application and then using them to derive model predictions”. Given the strong seasonality of snow-influenced discharge, several studies propose a step-wise calibration of snow hydrological models by calibrating model parameter groups on selected discharge periods (He et al., 2015; Hingray et al., 2010; Schaeffli et al., 2005). The diagnostic power of these approaches is, however, limited by the fact that they essentially rely on the minimization of model residuals (in particular the difference between observed and modelled discharge).

The identification of model parameters and model structures based on model residual selects a posteriori (after the model has been run) those simulations that are compatible with the observed data (in particular discharge), e.g. by minimizing some measure of the magnitude of the model residuals. The main difference between the various residual-based model identification methods is the a posteriori model selection procedure, which uses either a GLUE-based methodology (Generalized Likelihood Uncertainty Estimation, Beven and Binley, 1992; Beven and Binley, 2014), a single or multi-objective global optimization approach (e.g. Vrugt et al., 2003) or formal Bayesian parameter inference (e.g. Henn et al., 2015).

Any model identification approach selects a model as being an acceptable representation of the analyzed system if it satisfies some selection criteria over the calibration period and if the model performance does not too strongly decrease over the selected validation period. Common to most residual-based model selection methods is the absence of an objective, a priori evaluation of how well an acceptable model should behave over the observational period.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

To overcome this major drawback, Beven (2006) proposed to select acceptable models based on hydrologic signatures for which limits-of-acceptability can be specified before any model simulation is actually completed. This framework provides a straightforward way to select models by simply rejecting all models that do not fall into the a priori limits for all retained signatures. It furthermore assumes that all simulations that lie within all limits-of-acceptability are equifinal. Such a selection procedure namely avoids the problem of averaging several residual or signature-distance measures, which might typically result in selecting models that have a mediocre performance for each individual distance measure. Examples of application of the limits-of-acceptability approach are presented in the work of (Blazkova and Beven, 2009;Coxon et al., 2014;Liu et al., 2009;Westerberg et al., 2011;Winsemius et al., 2009).

The most widely used hydrologic signature for model identification is the flow-duration curve. Other signatures include the memory of the discharge (its autocorrelation, e.g. Toth, 2013;Winsemius et al., 2009), high or low flow frequency (Toth, 2013;Viglione et al., 2013) or the runoff ratio. Westerberg and McMillan (2015) provide an up-to-date overview of existing studies and of used signatures and propose a method to estimate their uncertainty.

In this paper, we propose a new way of extracting information from observed discharge in high mountainous environments and show how this information can be translated into snow hydrologic signatures to be used within a limits-of-acceptability approach. A new measure of the discriminatory power of these signatures is proposed to overcome the limitations related to subjective signature selection (Euser et al., 2013). The relevance of the proposed signature-based model selection framework is discussed for the Swiss Dischma case study, which is representative of many Alpine settings.

**2 Case study**

The case study is the Dischma river, located in the South-East of Switzerland near Davos (Figure 1). Its catchment has a size of 43.3 km<sup>2</sup> at the Kriegsmatten gauging station. The discharge time series for the years 1981 – 2012 are available from the Swiss Federal Office for the Environment. The catchment elevation ranges from 1668 m asl. to 3146 m asl, with a mean altitude of 2372 m asl.; around 2% of the catchment area is today covered by glaciers. The hillslopes are relatively steep with shallow soils, mainly covered with pasture (38%), coniferous forest (10%), bare soil (16%) and rock outcrops (24%) (Verbunt et al., 2003). The geology is crystalline composed of gneiss and amphibolites.

The discharge regime is of the so-called b-glacier type (Weingartner and Aschwanden, 1989), with a strong annual cycle resulting from accumulation and melt of snow. The monthly maximum streamflow occurs in early summer (June) and monthly minimum flow in January (Figure 2).

The observed air temperature data is taken from the meteorological station Davos (1594 m asl., Figure 1 of Supporting Material), 5.0 km downstream of the catchment outlet. Air temperature lapse rates are computed between this station and the nearby Weissfluhjoch station (2690 m asl., distance to outlet 8.6 km) and are around -0.5°C/100m. This results in a mean annual temperature at mean elevation of -0.2°C (1981-2012) but with a strong increasing trend over the study period (Figure 2). Albeit not located in the same valley, the precipitation recordings of the high elevation meteo station Weissfluhjoch (2690 m asl.) are used to simulate the discharge of this catchment. The station records around 1400 mm/year of precipitation, which is relatively low compared to other Alpine locations at the same altitude:

As illustrated in Figure 2, the annual totals can be assumed to be compatible with the streamflow recordings of the Dischmabach (1220 mm/year), whereas the recordings from the low elevation Davos station show too low annual totals. A detailed analysis showed that the pattern of precipitation occurrence does not differ significantly between the two meteorological stations at a daily scale. It can accordingly be assumed that, at a daily time step, the Weissfluhjoch time series is representative for the catchment.

It is noteworthy that at least over the most recent years, part of the discharge is probably due to net glacier ice melt. The mean evaporation in this catchment was of the order of magnitude of 300 mm/year for the period 1973 - 1992 (Menzel et al., 1999). There are no snow depth measurement stations directly located within this catchment.

## 2.1 Hydrological model set-up and Monte Carlo simulations

The discharge simulations are completed with the spatially-explicit hydrologic response model SEHR-ECHO (Schaepli et al., 2014). This model simulates the hydrologic response of 23 subcatchments identified from the river network with TauDEM Version 5 (Tarboton, 1997), which is freely available from <http://hydrology.usu.edu/taudem/taudem5/>. The network and the subbasins are shown in Figure 1. The temperature time series for each subcatchment is obtained with a constant air temperature lapse rate. Given the absence of detailed information on the spatial distribution of precipitation, the observed data from Weissfluhjoch station is not interpolated to the subcatchments. The potential evaporation is evaluated with sunshine data from the Weissfluhjoch station and the Priestley-Taylor method (Maidment, 1993).

The core of the precipitation-runoff simulation for each spatial unit (subcatchment) is a simple snow accumulation and melt routine based on a degree-day approach and a step transition between snow- and rainfall at 1°C. For partially ice-covered subcatchments, ice melt starts as soon as the entire snowpack has melted away and is also simulated with a simple degree-day approach (which is deemed reliable enough given the small extend of the ice-covered area).

Evaporation is computed as direct re-evaporation of intercepted precipitation if there is enough potential evaporation. Transpiration is computed as a function of soil moisture. Sublimation from the snowpack or from the glacier surface is neglected.

The transformation of melt water and rainfall into runoff at the subcatchment scale (before it enters a stream) is completed with a simple soil moisture accounting routine and two subsurface flows for the slow and the fast flow component. Routing along the network is obtained via convolution with the river network. For the daily time step considered here, routing has however no effect.

To illustrate how the presented model selection framework can be used for hypothesis testing, two different spatial model set-ups are used: i) the basic set-up with 23 subcatchments, ii) a set-up with five elevation bands per subcatchment (115 spatial units in total). For this second set-up, the subcatchment state variables (snow storage, soil moisture etc) and flux variables (ice melt flux, snowpack outflux, fast and slow subsurface flows) are obtained by area-averaging the state variables and the flux variables of the corresponding elevation bands.

These two spatial set-ups are interesting because (2014) the subcatchment set-up gives a better description of the origin and the flow paths of water, whereas the elevation band set-up captures better the altitudinal distribution of snow accumulation and melt (Schaepli et al., 2014). Based on simple minimization of discharge residuals, it is, however, impossible to show that the elevation set-up gives significantly better results at the daily time step.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

In the two set-ups used here, the model has a total of 12 parameters to calibrate (Table 1). Parameters that are fixed are the interception threshold (1 mm), the snow refreezing factor (0.2 times the degree-day factor), snow retention capacity (0.05), soil porosity (0.4) and the maximum soil infiltration capacity (set to infinity, to switch off the Hortonian overland flow). All the results presented here are obtained with a total of 100'000 Monte Carlo simulations obtained by uniform random draws of parameter sets in the prior parameter ranges of Table 1. This number of simulations is a heuristic choice, resulting from repeated draws of parameter ensembles that showed that model performance does only very slightly increase beyond a draw of around 40'000 samples. The retained parameter ensemble does not contain parameters that do not respect the constraints indicated in Table 1. To initialize the seasonal snowpack and the soil moisture store, the first two years of each simulation are discarded for model performance analysis and for signature computation.

**3 Method**

**3.1 Snow information extraction from hydrometeorological data**

The annual discharge cycle of snow-dominated catchments is due to the seasonal storage of water in the snowpack, which results in a long recession in autumn and throughout the winter and in a period of increasing snowmelt input overlain by rainfall events. In Alpine environments, two of the three dominant energy sources for snowmelt are long wave radiation and sensible heat, which are strongly influenced by air temperature (Ohmura, 2001). Given that air temperature also influences the aggregation state of precipitation (snow versus rain), air temperature can be considered as being the dominant driver of the annual discharge cycle resulting from snow accumulation and melt. The timing of this discharge cycle is fairly constant from year to year because air temperature follows a relatively invariant seasonal cycle. Accordingly, a simple but effective way to make these seasons visible is to compute the interannual mean of the discharge on each calendar day (Hingray et al., 2010) (Figure 3). Plotting this interannual discharge regime against the air temperature regime (Figure 3) nicely shows how discharge almost linearly increases with air temperature in late spring. For the Dischma, this slope varies between 1 and 6 °C/(m<sup>3</sup>/s) with the exception of the period centred on the extremely snow-rich winter 1999/2000 (Wiesinger, 2006), where the slope is 10°C/(m<sup>3</sup>/s). It is noteworthy that this slope cannot directly be translated into snow melt factors since the slope results from a combination of snowmelt and an increase of the snowmelt-affected area. Once the snowpack has sufficiently melted away, any further increase in temperature cannot result in a further discharge increase (turning point in the discharge-temperature cycle). In late summer, air temperature also starts decreasing, leading ultimately in Autumn to an almost linear relationship between discharge decrease and air temperature due to the progressive increase of the non-responsive catchment area (Schaepli et al., 2013). This discharge-air temperature cycle is typical for snow-dominated catchments. For highly glacier-covered catchments, the discharge increase with air temperature in spring/early summer has a very similar slope to the discharge decrease with air temperature in late summer/autumn (see Figure 2 of the Supporting Material). This essentially reflects the increase and decrease of the responsive area as a function of air temperature. For catchments

with low glacier cover, the spring and autumn slopes differ strongly because the second is influenced also by the catchment recession behaviour.

Further information on the snowpack accumulated in the catchment can be extract by considering the difference between cumulated catchment input (precipitation) and cumulated catchment outflow following the mass curve technique (MCT) proposed by Gao et al. (2014) to estimate the root zone depth. Knowing that the precipitation exceeding discharge during winter is essentially stored as snow and that the amount of discharge exceeding precipitation during spring has to come from snow storage, we can obtain two estimates of the amount of catchment-wide snow storage per year, one from winter discharge and one from summer discharge (see Figure 3 Supporting Material). Comparison of these two estimates from observed discharge and from simulated discharge can provide further insights into the model's ability to reproduce the observed hydrologic response.

For the case study at hand, the discharge gauging station at the outlet records also water temperature, which potentially represents an interesting source of information on dominant hydrological processes. In such a snow-dominated catchment, the different runoff components (glacier melt, direct surface runoff from rainfall or from snowpack outflow, slow and fast subsurface runoff) have in fact very different temperature signatures (Comola et al., 2015a). The Dischmabach gauging station is, however, located relatively far downstream on the main river channel, such that the water temperature is dominated by in-stream heat exchange with the atmosphere during a large part of the year (even if there is a slight distributional difference for low water temperatures during days with discharge increase compared to recession days, Figure 4 Supporting Material). This is in fact the case for the majority of Swiss water temperature gauging stations (Gallice et al., 2015; Marco and Sebastiano, 2015). The strong atmospheric influence is highlighted in the plot of the water temperature against the air temperature regime in Figure 4, which shows only a very small hysteresis effect between the two signals.

Accordingly, from a hydrological modelling perspective, river water temperature measurements essentially reflect the same information as contained in direct air temperature observations.

### 3.2 Snow hydrological signatures

The slopes of the above discharge-temperature regime cycle over the identified seasons form a first set of potential signatures candidates. The timing of the transitions between seasons is not used as a signature since its value would too strongly depend on algorithmic choices to identify the transitions.

The value of these slopes for model selection for each of the seasons is:

- *Early spring, early summer, winter*: these seasons cannot be used as signatures due to the non-linear link between discharge and temperature
- *Late spring*: the model reproduces the spatially-integrated link between temperature forcing and snowmelt, integrating namely the spatial distribution of the responsive area and the relationship between air temperature and snow melt
- *Summer and late summer*: the model reproduces the relation between air temperature and discharge in presence of transpiration and evaporation, with minor melt inputs from very high elevation snowpacks or from glaciers
- *Autumn*: the model reproduces the temporal decrease of the responsive area and the recession behaviour of the catchment.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

To further characterize the discharge regime and in particular the winter recession, the accumulation and melting of snow, we propose to use the timing of the maximum value of the discharge regime (in days) and the relative ratio between the maximum and the minimum value of the discharge regime according to:

$$r_Q = \frac{\max(\bar{Q}) - \min(\bar{Q})}{\max(\bar{Q})} \tag{2}$$

where  $\bar{Q}$  is the interannual discharge regime as discussed earlier. This ratio should be relatively stable since periods with strong snow accumulation result in low  $\min(\bar{Q})$  and high  $\max(\bar{Q})$ , whereas periods with low snow accumulation result in a small difference between  $\min(\bar{Q})$  and  $\max(\bar{Q})$  and in a lower  $\max(\bar{Q})$ .

Note that this ratio is similar to the Pardé coefficient range used by Viglione et al. (2013); the Pardé coefficients are, however, defined over monthly discharge, which represents an unnecessary dimensional reduction of the daily data.

Furthermore, a set of classically used discharge signatures (Viglione et al., 2013;Toth, 2013) is tested for illustrative purposes and for a concise description of the hydrological behaviour over the observation period, including the mean discharge, the slope of the flow-duration curve between percentile 30% and 70%, the high flow (percentile 95%) and the low flow (percentile 5%), the lag-1 autocorrelation, the time lag at which discharge autocorrelation drops below 1/e (a measure of the memory) and the coefficient of variation over the snow melt months (mean discharged divided by standard deviation). All signatures used in this paper are summarized in Table 2.

The snow accumulation estimated from observed and simulated discharge will serve as a further plausibility check of the involved hydrometeorological data.

**3.3 Limits-of-acceptability**

The definition of limits-of-acceptability for a selected signature should account for the uncertainty about the value of this signature (Blazkova and Beven, 2009). McMillan et al. (2012) provide an in-depth discussion of the observational sources of signature uncertainty. Westerberg and McMillan (2015) propose to address these uncertainties by explicitly quantifying observational uncertainties within a Monte Carlo simulation framework.

Here, we propose to adopt a much simpler method, which consists of splitting the available discharge time series into several subsamples as it is classically done e.g. for flow-duration curve uncertainty analysis (Vogel and Fennessey, 1994;Viglione et al., 2013).

Such an extension of the single observational sample into a set of samples has, however, an important limitation: In case of overlapping samples, the signature values estimated from each sample are not independent, which would impact the statistical interpretation of the signature distributions but does not affect the definition of limits-of-acceptability.

For the definition of the limits-of-acceptability based on these reference values, we distinguish two cases here: i) For signatures that are computed directly on the observed daily values, the limits-of-acceptability are the limits of the observed values. For the high flow signature for example, the limits-of-acceptability correspond to the lowest and the highest 95% flow quantile observed for all reference samples. ii) For signature values that are obtained from some averaging of the observed sample, we do not consider the lowest and the



highest reference value if these extreme values are obtained from averaging over shorter samples than for the simulated signature (as is the case for the results in presented in Section 4). This is for example the case for the spring  $T$ - $Q$  slope.

### 3.4 Discriminatory power of signatures

We propose here to assessed the discriminatory power of a given signature using two criteria: First, the average width of the simulation range spanned by acceptable simulations:

$$R_i = \frac{1}{N} \sum_{t=1}^N \frac{[q_{sim,max}(t|s_i) - q_{sim,min}(t|s_i)]}{q_{sim,50}(t)}, \quad (2)$$

where  $q_{sim,min}(t|s_i)$ ,  $q_{sim,max}(t|s_i)$  and  $q_{sim,50}(t|s_i)$  are, at time step  $t$ , the minimum value, the maximum value and the median value of all  $N$  acceptable simulations under signature  $s_i$ .

The second criterion  $\varphi_i$  measures whether under signature  $s_i$  the observations are uniformly distributed within the simulation range rather than e.g. always close to the limits of the simulation range. This is a necessary requirement for good prediction ranges (Laio and Tamea, 2007; Thyer et al., 2009). Following the reasoning of Laio and Tamea (2007) we propose to quantify  $\varphi_i$  as the maximum deviation of the empirical cumulative distribution function from the uniform cumulative distribution. Furthermore, we evaluate this maximum deviation per season to measure whether the observations are uniformly distributed within the simulation range for each season  $\mathbf{T}_j = \{\text{DJF}, \text{MAM}, \text{JJA}, \text{SON}\}$ :

$$\varphi_{i,j} = \max |P_{i,j}(q_{obs}(t|t \in \mathbf{T}_j)) - F_U| \quad (3)$$

$$\varphi_i = \max(\varphi_{i,j}) \quad (4)$$

where  $F_U$  is the cumulative uniform distribution  $U(0,1)$  and  $P_{i,j}(q_{obs}(t|t \in \mathbf{T}_j))$  is the empirical distribution function of  $q_{obs}(t)$  within the simulation range under signature  $s_i$ . This measure  $\varphi_i$  is much less sensitive to observational errors than for example the range of observed values covered by the simulation range.

A signature with high discriminatory power should result in a very narrow simulation range ( $R_i \ll 1$ ) with uniformly distributed observations within this range for all seasons ( $\varphi_i$  close to zero). We thus use a criteria for the discriminatory power of a signature  $D_i$  computed as

$$D_i = 1 - \varphi_i R_i = 1 - \frac{\varphi_i}{N} \sum_{t=1}^N \frac{[q_{sim,max}(t|s_i) - q_{sim,min}(t|s_i)]}{q_{sim,50}(t)}, \quad (5)$$

which should be close to 1.

### 3.5 Additional model performance measures

For model performance evaluation, the following criteria are used in addition to the above signatures: the Nash-Sutcliffe criterion (Nash and Sutcliffe, 1970), called here simply Nash value, the Nash value computed on the log-transformed values (called Nash-log) and the benchmark efficiency criterion ( $B_E$ ) suggested by Schaefli and Gupta (2007), where the reference model is the interannual mean-discharge on each calendar day rather than the mean discharge over the entire period:

$$B_E = 1 - \frac{\sum_{t=1}^N [q_{obs}(t) - q_{sim}(t)]^2}{\sum_{t=1}^N [q_{obs}(t) - q_b(t)]^2} \quad (5)$$

where  $q_{obs}(t)$  is the observed discharge at time step  $t$ ,  $q_{sim}(t)$  is the simulated discharge and  $q_b(t)$  is the benchmark model discharge at time step  $t$  obtained as the interannual mean on the calendar day corresponding to  $t$ . As discussed by Schaepli and Gupta (2007), this benchmark efficiency measures whether the hydrologic model explains more of the observed variability than what is already contained in the seasonality of the climate.

## 4 Results and discussion

### 4.1 Model calibration on Nash values

As a benchmark for the results presented hereafter, Figure 5 shows the best discharge simulations obtained under the two studied model set-ups, calibrated either on the snow-rich period 1983-1992 or on the relatively warm period 2000-2009, which includes the heat wave summer 2003. The different calibrations yield relatively similar parameter values for snowmelt and the soil moisture module but yield in particular rather different time scales for the fast and the slow subsurface flow components. As a result, the discharge simulations are clearly very different for the years shown in Figure 1; all simulations do a rather poor job for the year 2000 (which followed an extremely snow reach winter) and for the summer 2003. Even the calibration on the period including the summer 2003 does not give good results here. In fact, the model performance shows an extreme performance drop after the year 2000 (see also Section 4.3), which points towards a major issue in either i) the model selection, ii) the model structure for warm and extremely warm years, iii) the used input data for this period or iv) the observed discharge. Understanding and overcoming this performance drop was thus a strong driver for the research presented here.

### 4.2 Snow accumulation estimated from discharge and precipitation

The precipitation accumulation during the winter low flow period estimated from observed precipitation and discharge represents a simple but potentially interesting signature to compare the model behaviour integrated over the winter period to the observed discharge.

The mass curve technique (MCT) is used to estimate catchment-scale accumulated water either from winter precipitation exceedance or from summer flow exceedance. The two estimates show a certain scatter for observed discharge as well as for an example of simulated discharge (Figure 6, left), which is partly related to the assumption of a constant slope of precipitation. Overall, the result for observed discharge suggests that the used precipitation is not incompatible with the observed discharge.

For the model simulations, MCT yields estimates from the winter flow that are very close to the actually simulated maximum winter accumulation (Figure 6, center), which shows that the method is promising for snowpack characterization from winter flow; the snowpack estimations from summer flow are underestimated, which is to be expected given that part of the water leaves the system via evaporation and transpiration. The simulated maximum winter accumulation is also in very good agreement with the snowpack estimations obtained with MCT from the observed flow (Figure 6, right) Plotting the snowpack estimates from observed

and simulated winter flow against each other underlines the coherence of the results (Figure 6, right). Albeit being a strong hint for a good spatially and temporally integrated simulation, these results do, however, not necessarily mean that the actual snowpack is well captured by the model. This would in particular not be the case if water is stored in some other form.

### 4.3 Time variability of signatures

The discharge observation period covers a certain number of exceptional years, starting with the overall snow-rich period in the 80ies (resulting in an advancing period for many Swiss glaciers, e.g. Bauder, 2015), the exceptional snow-rich winter 1999/2000, the heat wave summer 2003 (Zappa and Kan, 2007), a period of higher precipitation recordings (1999-2004, Figure 2) and the overall warm period since 2000 (see also an analysis of climatic conditions for glaciers for four Swiss study sites in Huss et al., 2009).

Accordingly, the signatures analyzed here show a strong time variability, which is not well captured by the model if calibrated on the Nash value. This is illustrated in Figure 7 that shows selected signature values computed over periods of 5 years with a moving window (overlapping periods). The Nash-optimal simulations result in an important over-estimation of the mean discharge in the precipitation-rich years, in an overall too early maximum discharge, too low flow-duration curve slope for mid-range discharges, too high low flows for two simulations, too low high flows for the recent warm years and a too low coefficient of variation of summer flows. The autocorrelation signatures show an important difference between the Nash-optimal simulations selected on period 1 (1983-1992) and the ones selected on period 2 (2002-2010): The signatures show an overall overestimation of the time scale of the autocorrelation decrease for the selection on period 2, an overestimation of the autocorrelation lag-1 for the subcatchment and for the elevation set-up for the two different periods and a complete mismatch of the autocorrelation lag-1 for the subcatchment set-up selected on period 1. This last result can be explained by the fact that the Nash-optimal simulation over period 1 leads to too frequent, too strong high flows over the remaining warmer years.

These time-variable differences between signatures estimated from simulated and from observed discharge are reflected in the strongly time-varying Nash performance (Figure 7, last subplot), with a remarkable drop of the Nash value around after 2000 for all set-ups but in particular for the simulations calibrated over 1983-1992. The benchmark efficiency values  $B_E < 0$  (same graph with right-hand scale) clearly show that the simulation does not explain any observed variability beyond the annual cycle. The Nash-log values (not shown) are high (above 0.85) for all simulations but also with a pronounced drop after 2000.

To summarize, it is clear from this analysis of the performance variability through time that a simple Nash-based parameter selection is likely to not result in a robust model calibration for this case study.

### 4.3 Discriminatory power of signatures

The simulation ranges obtained under the different signatures studied in this paper show important differences in their width over the different seasons and in the way they include the observed discharge values. An example is given in

Figure 10 for the timing of the annual maximum  $Q_{\max}^*$ . The discriminatory power  $D_i$  of the studied signatures (Figure 8) shows that only a few signatures have a positive value, i.e. only a few signatures lead to a relatively narrow simulation range that reasonably covers the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

observed values (see also the components of  $D_i$  split according to the seasons in the Supporting Material, Figure 5).

In the following, we only consider these signatures with positive  $D_i$  for further model selection, i.e. the timing of the maximum regime discharge (max. timing), the coefficient of variation during the melt season (cv melt), the temperature-discharge regime slope during spring (T-Q spring slope) and the high flows. This set of signatures is completed with the lag-1 autocorrelation as a minimum control on the temporal structure of the model. All these signatures essentially focus on high flows during spring and summer. It will be particularly interesting to see whether models that respect these signatures also perform well during autumn and winter.

The selected signatures apply a strong selection on the model parameters as illustrated in the dotty-plots (Beven, 2006) for the degree-day factor of snow (Figure 9). For this example, the parameter selection resulting from the individual signatures is visible even if we consider the model performance projection onto a single parameter axis (comparable to considering the marginal distribution in a formal parameter inference approach). For most other parameters, this effect is not visible in the single parameter space: many signatures seem to not exert any selection effect on single parameters (Supporting Material, Figure 6), despite of their discriminatory power on the model prediction range. This absence of a selection effect of an individual signature on a single parameter does, however, not mean that the parameter is not identifiable under this signature; this might also result from parameter correlations, as nicely demonstrated by Bardossy (2007).

**4.4 Signature-based model improvement and model selection**

Having selected a range of signatures that have useful discriminatory power for the selected model and the hydroclimatic environment, these signatures can be used for model improvement by analyzing the acceptable simulation range for exceptional years. For the Alpine case studies, the heat wave year 2003 is an ideal candidate for this.

Figure 10 top row shows the range of simulations spanned by models that are within the limits-of-acceptability for the max. timing. The simulation range is clearly reduced to a space that is far off the observations in summer 2003. Introducing five elevation bands per subcatchment and a time-variable lapse rate partially solves this issue (

Figure 10 bottom row).

With this new model structure, only 26 out of 100'000 parameter sets lead to simulations that are acceptable for all the five retained signatures simultaneously (Figure 11). The resulting simulation range corresponds reasonably well to the observed discharge for the two extreme years. More interestingly, these simulations capture the other signatures, not retained for model selection amazingly well (Figure 12). This simply expresses the fact that for snow-dominated systems, capturing the essential features of accumulation and release of snow captures the key dynamics.

It is noteworthy that some simulations that are acceptable for the five retained signatures show a clearly lower Nash performance than the others(Figure 12, last plot), which would require some detailed analysis to understand why they nevertheless yield good signature values.

#### 4.5 So what is going on here after the year 2000?

No simple model structure modification nor any sophisticated signature-based model selection procedure can improve the model performance over the period 2000-2005. A check of the used meteorological time series recorded at two rather different locations (Davos station at 1594 m asl. in a relatively flat area in the village close to the catchment outlet and the exposed mountain top Weissfluhjochstation at 2690 m asl.) shows that a problem with these stations can be excluded since they both behave very similarly over the considered period. None of the stations has been modified or has experienced significant problems (T. Konzelmann, MeteoSwiss, personal communication Nov. 2015).

An analysis of yearly double mass curves of the discharge of the Dischma river and of a downstream gauge (Landwasser, catchment area 183 km<sup>2</sup>) suggests that the two discharge time series do not scale well for some years of the critical period, but not for all of them (Schaepli et al., 2016).

A regional check with discharge data from catchments with similar settings shows, however, a remarkable result: The simulated annual discharge anomalies (with respect to a longer term mean) for the Dischma catchment (with time variable lapse rates) closely follow the trend of the nearby rivers (Figure 13), while the observed data completely drifts away. In addition, all published simulations with similar models (with PREVAH and HBV by Orth et al., 2015; with HBV by Jenicek et al., 2016) show a comparable deviation from the observed discharge data over the period 2000-2005 (Schaepli et al., 2016).

A closer look on the discharge data, reveals, in addition, that the hourly data shows inexplicable sub-daily fluctuations for the year 2004 (see Supporting Material, Figure 7), which disappear in the year 2005. The managers of the station (Swiss Federal Office for the Environment) have not noticed this before but mentioned on request that the old analogue limnigraph was replaced on 27 Oct 2003 with a pressure sensor and that this sensor was replaced again in November 2006 and in September 2011 (A. Kohler, personal communication, 12 October 2015). No further details are available at the moment.

All these hints strongly suggest that the observed data requires a much closer inspection before further modeling work can be done here. Discussions with hydrologists from several Swiss research groups are ongoing (Schaepli et al., 2016).

## 5 Discussion

The retained set of signatures contains regime-based signatures (spring T-Q slope, max. timing) as well as signatures computed from the discharge time series at a daily time step (cv melt, AR1, high flow). As discussed by Coxon et al. (2014), regime-based signatures are particularly important for catchments where the discharge is dominated by seasonal and climatic features rather than individual rainfall events, which is typically the case for snow-dominated environments.

All the selected signatures, except the signature on high flows, result from average properties and are thus relatively robust against non-systematic observational errors and against disinformative data sequences (Beven and Westerberg, 2011); this makes them particularly useful for model identification in presence of observational errors (McMillan et al., 2012; Westerberg and McMillan, 2015). Furthermore, most of the proposed signatures can be computed with no computational choices (except the T-Q slopes), which are a non-negligible source of signature uncertainty (Westerberg and McMillan, 2015). In addition, the proposed



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

regime-based signatures might be relatively constant across similar catchments and they present thus a high potential for signature regionalization (Ganora et al., 2009;Yadav et al., 2007).

Other studies in snow-dominated environments used signatures that are similar to the set of regime-based signatures investigated here (e.g. Kelleher et al., 2015) but they were generally based on annual values. The model constraining effect or the discriminatory power of such apparently similar annual signatures might be relatively different. In addition, the value of similar signatures for model identification can only be compared if they are applied to models that have similar functional units, i.e. whose model components simulate dominant processes with similar state and flux variables.

The presented detailed study of a single catchment with a range of tailor-made signatures underlines the value of the limits-of-acceptability approach developed by Keith Beven for model identification and hydrological system analysis in presence of error-prone observational data. This model selection framework allows to combine our a priori knowledge about the model behaviour with a priori knowledge about the catchment dynamics (Wagener and Montanari, 2011) without condensing this knowledge into some model performance measure. Such an explicit analysis of the model behaviour with respect to different signatures is particularly important for snow-dominated environments, where compensation effects between water released from different stores are potentially high.

An important aspect of the presented study is the analysis of the constraining effect of the signatures in the model prediction space rather than in the parameter space as in many of Beven's papers. This shift to the prediction space is particularly important for snow-dominated environments where model parameters are strongly correlated because the filling and emptying of different water stores (soil, snow, ice) depend on each other at different spatio-temporal scales. In general, the proposed analysis of the discriminant power of the signatures in the model prediction domain represents a step towards the development of more informative signatures (Wagener and Montanari, 2011).

An important open question is in as far the proposed signatures and the limits-of-acceptability approach in general can be used for ungauged catchments, since this would require the definition of regional limits-of-acceptability in absence of local discharge observations.

**6 Conclusion**

In snow-dominated environments, the dominant hydrological processes have a relatively stable imprint on the discharge regime, due to the pronounced annual storage and release of water. The extraction of process information directly from the discharge regime has nevertheless received relatively little attention in the past. As illustrated here based on a Swiss case study (the Dischma river), namely the average timing of annual maximum flow and the relationship between the air temperature regime and the discharge regime have a high potential for model selection based on theory and on physical process understanding (Clark et al., 2016) rather than based on residual analysis. Future work will show how the relationship between these two regimes can be used for catchment similarity assessment. The proposed method to estimate winter snow accumulation directly from observed discharge represents, in addition, a new simple method for plausibility checks of the available hydrometeorological data.

The selected case study illustrates nicely how model selection with a priori limits-of-acceptability can decisively contribute to learn from the data to improve the model structure, e.g. by inspecting the acceptable simulation range for extreme years. Even models selected

over a period of possibly wrong discharge observations have been shown to still lead to acceptable system representations over the remaining period. The potential of the presented signatures for model selection in the presence of non-concomittant time series remains to be analysed (De Vleeschouwer and Pauwels, 2013; Montanari and Toth, 2007; Winsemius et al., 2008).

Simple snow hydrological models have a long-standing tradition in hydrologic prediction. Given the impressive spatial heterogeneity of the involved processes, the good performance of simple models continues however to trigger suspicion. There even seems to be a general agreement that data-intense, grid-based energy-balance models are in principle preferable for catchment-scale prediction - despite of all discussions and scientific evidence on the limitations of complex physics-based models that we owe to the work of Keith Beven since his (1989) paper.

The snow hydrological signatures proposed here within a limits-of-acceptability approach for model selection has the potential to shed a new light on this ‘simple versus complex models’ debate for Alpine environments. If a simple model that yields acceptable regime-based flow signatures also reproduces reasonably well the daily flow patterns, this is a strong hint that the model simplifications are justified at the considered spatio-temporal scale. A complex snow-model without calibration of accumulation and of melt parametrizations might, in exchange, only be an acceptable representation of the system behaviour if it is able to capture the a priori limits-of-acceptability of some key signatures. From this perspective, Keith Beven’s limits-of-acceptability approach has not yet unraveled its full potential for model development in snow-dominated environments.

## Acknowledgement

The research of the first author has been supported by the Swiss Competence Center on Energy Research – Supply of Electricity, with the support of the Swiss Commission for Technology and Innovation. The meteorological data were obtained from MeteoSwiss and the discharge and water temperature data from the Swiss Federal Office for the Environment.

## References

- Bardossy, A.: Calibration of hydrological model parameters for ungauged catchments, *Hydrology and Earth System Sciences*, 11, 703–710, 10.5194/hess-11-703-2007, 2007.
- Bauder, A.: The Swiss Glaciers 2009/10 and 2010/11, *Glaciological Report No 131/132*, available at: [http://glaciology.ethz.ch/messnetz/downloadPubs/glrep\\_131-132.pdf](http://glaciology.ethz.ch/messnetz/downloadPubs/glrep_131-132.pdf), Zürich, 119, 2015.
- Beven, K.: Changing ideas in hydrology—the case of physically-based models, *Journal of Hydrology*, 105, 157–172, 1989.
- Beven, K., and Westerberg, I.: On red herrings and real herrings: disinformation and information in hydrological inference, *Hydrological Processes*, n/a-n/a, 10.1002/hyp.7963, 2011.
- Beven, K., and Binley, A.: GLUE: 20 years on, *Hydrological Processes*, 28, 5897-5918, 10.1002/hyp.10082, 2014.
- Beven, K. J., and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, *Hydrological Processes*, 6, 279-298, 10.1002/hyp.3360060305, 1992.
- Beven, K. J.: Prophecy, Reality and Uncertainty in Distributed Hydrological Modeling, *Advances in Water Resources*, 16, 41-51, 10.1016/0309-1708(93)90028-E, 1993.
- Beven, K. J.: A manifesto for the equifinality thesis, *Journal of Hydrology*, 320, 18-36, 10.1016/j.jhydrol.2005.07.007, 2006.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Blazkova, S., and Beven, K.: A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, *Water Resour. Res.*, 45, W00B16, 10.1029/2007wr006726, 2009.

Clark, M. P., Schaefli, B., Schymanski, S. J., Samaniego, L., Luce, C. H., Jackson, B. M., Freer, J. E., Arnold, J. R., Dan Moore, R., Istanbuluoglu, E., and Ceola, S.: Improving the theoretical underpinnings of process-based hydrologic models, *Water Resources Research*, n/a-n/a, 10.1002/2015wr017910, 2016.

Comola, F., Schaefli, B., Rinaldo, A., and Lehning, M.: Thermodynamics in the hydrologic response: Travel time formulation and application to Alpine catchments, *Water Resources Research*, n/a-n/a, 10.1002/2014wr016228, 2015a.

Comola, F., Schaefli, B., Ronco, P. D., Botter, G., Bavay, M., Rinaldo, A., and Lehning, M.: Scale-dependent effects of solar radiation patterns on the snow-dominated hydrologic response, *Geophysical Research Letters*, 42, 2015GL064075, 10.1002/2015gl064075, 2015b.

Coxon, G., Freer, J., Wagener, T., Odoni, N. A., and Clark, M.: Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments, *Hydrological Processes*, 28, 6135-6150, 10.1002/hyp.10096, 2014.

De Vleeschouwer, N., and Pauwels, V. R. N.: Assessment of the indirect calibration of a rainfall-runoff model for ungauged catchments in Flanders, *Hydrol. Earth Syst. Sci.*, 17, 2001-2016, 10.5194/hess-17-2001-2013, 2013.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures, *Hydrol. Earth Syst. Sci.*, 17, 1893-1912, 10.5194/hess-17-1893-2013, 2013.

Finger, D., Pellicciotti, F., Konz, M., Rimkus, S., and Burlando, P.: The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model, *Water Resour. Res.*, 47, W07519, 10.1029/2010wr009824, 2011.

Gallice, A., Schaefli, B., Lehning, M., Parlange, M. B., and Huwald, H.: Stream temperature prediction in ungauged basins: review of recent approaches and description of a new physics-derived statistical model, *Hydrol. Earth Syst. Sci.*, 19, 3727-3753, 10.5194/hess-19-3727-2015, 2015.

Ganora, D., Claps, P., Laio, F., and Viglione, A.: An approach to estimate nonparametric flow duration curves in ungauged basins *Water Resources Research*, 45, W10418,, 10.1029/2008WR007472 2009.

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., and Savenije, H. H. G.: Climate controls how ecosystems size the root zone storage capacity at catchment scale, *Geophysical Research Letters*, 41, 7916-7923, 10.1002/2014gl061668, 2014.

Gupta, H. V., Wagener, T., and Liu, Y. Q.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrological Processes*, 22, 3802-3813, 10.1002/hyp.6989, 2008.

He, Z. H., Parajka, J., Tian, F. Q., and Blöschl, G.: Estimating degree-day factors from MODIS for snowmelt runoff modeling, *Hydrol. Earth Syst. Sci.*, 18, 4773-4789, 10.5194/hess-18-4773-2014, 2014.

He, Z. H., Tian, F. Q., Gupta, H. V., Hu, H. C., and Hu, H. P.: Diagnostic calibration of a hydrological model in a mountain area by hydrograph partitioning, *Hydrology and Earth System Sciences*, 19, 1807-1826, 10.5194/hess-19-1807-2015, 2015.

Henn, B., Clark, M. P., Kavetski, D., and Lundquist, J. D.: Estimating mountain basin-mean precipitation from streamflow using Bayesian inference, *Water Resources Research*, n/a-n/a, 10.1002/2014wr016736, 2015.

Hingray, B., Schaefli, B., Mezghani, A., and Hamdi, Y.: Signature-based model calibration for hydrologic prediction in mesoscale Alpine catchments, *Hydrological Sciences Journal*, 55, 1002-1016, 10.1080/02626667.2010.505572, 2010.

Huss, M., Funk, M., and Ohmura, A.: Strong Alpine glacier melt in the 1940s due to enhanced solar radiation, *Geophysical Research Letters*, 36, L23501, 10.1029/2009GL040789, 2009.

- Jenicek, M., Seibert, J., Zappa, M., Staudinger, M., and Jonas, T.: Importance of maximum snow accumulation for summer low flows in humid catchments, *Hydrol. Earth Syst. Sci.*, 20, 859-874, 2016.
- Kelleher, C., Wagener, T., and McGlynn, B.: Model-based analysis of the influence of catchment properties on hydrologic partitioning across five mountain headwater subcatchments, *Water Resources Research*, 51, 4109-4136, 10.1002/2014wr016147, 2015.
- Laio, F., and Tamea, S.: Verification tools for probabilistic forecasts of continuous hydrological variables, *Hydrol. Earth Syst. Sci.*, 11, 1267-1277, 10.5194/hess-11-1267-2007, 2007.
- Liu, Y., Freer, J., Beven, K., and Matgen, P.: Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error, *Journal of Hydrology*, 367, 93-103, 10.1016/j.jhydrol.2009.01.016, 2009.
- Maidment, D. R.: *Handbook of Hydrology*, in, McGraw-Hill, New York, 1993.
- Marco, T., and Sebastiano, P.: A hybrid model for river water temperature as a function of air temperature and discharge, *Environmental Research Letters*, 10, 114011, 2015.
- McMillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality, *Hydrological Processes*, 26, 4078-4111, 2012.
- Menzel, L., Lang, H., and Rohmann, M.: Mean annual actual evaporation 1973 - 1992, in: *Hydrological Atlas of Switzerland*, Plate 4.1, Service Hydrologique et Géologique National, Berne, 1999.
- Montanari, A., and Toth, E.: Calibration of hydrological models in the spectral domain: an opportunity for ungauged basins?, *Water Resources Research*, 43, W05434, 10.1029/2006WR005184, 2007.
- Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models. Part I, a discussion of principles, *Journal of Hydrology*, 10, 282-290, 10.1016/0022-1694(70)90255-6 1970.
- Ohmura, A.: Physical basis for the temperature-based melt-index method, *J. Appl. Meteorol.*, 40, 753-761, 2001.
- Orth, R., Staudinger, M., Seneviratne, S. I., Seibert, J., and Zappa, M.: Does model performance improve with complexity? A case study with three hydrological models, *Journal of Hydrology*, 523, 147-159, 10.1016/j.jhydrol.2015.01.044, 2015.
- Parajka, J., and Blöschl, G.: The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models, *Journal of Hydrology*, 358, 240-258, 10.1016/j.jhydrol.2008.06.006, 2008.
- Schaeffli, B., Hingray, B., Niggli, M., and Musy, A.: A conceptual glacio-hydrological model for high mountainous catchments, *Hydrology and Earth System Sciences*, 9, 95 - 109, 10.5194/hessd-2-73-2005, 2005.
- Schaeffli, B., and Gupta, H.: Do Nash values have value?, *Hydrological Processes*, 21, 2075-2080, 10.1002/hyp.6825, 2007.
- Schaeffli, B., Rinaldo, A., and Botter, G.: Analytic probability distributions for snow-dominated streamflow, *Water Resources Research*, 49, 1-13, 10.1002/2012WR020234, 2013.
- Schaeffli, B., Nicótina, L., Imfeld, C., Da Ronco, P., Bertuzzo, E., and Rinaldo, A.: SEHR-ECHO v1.0: a Spatially Explicit Hydrologic Response model for ecohydrologic applications, *Geosci. Model Dev.*, 7, 2733-2746, 10.5194/gmd-7-2733-2014, 2014.
- Schaeffli, B., Ceperley, N., Brauchli, T., Lehning, M., Jonas, T., and Zappa, M.: The Dischma river mystery: why does my snow hydrological model not work here ? *Geophysical Research Abstracts* Vol. 18, EGU2016-1624, 2016, EGU General Assembly 2016, Vienna, 2016.
- SwissTopo: DHM25- The digital height model of Switzerland, Wabern, Switzerland, 2005.
- SwissTopo: Vector25 - The digital landscape model of Switzerland, Wabern, Switzerland, 2008.
- Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, *Water Resources Research*, 33, 309-319, 10.1029/96wr03137, 1997.
- Thyer, M., Benjamin Renard, B., Dmitri Kavetski, D., George Kuczera, G., Stewart Franks, S., and Sri Srikanthan, S.: Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, *Water Resources Research*, 45, W00B14, 10.1029/2008WR006825, 2009.
- Toth, E.: Catchment classification based on characterisation of streamflow and precipitation time series, *Hydrol. Earth Syst. Sci.*, 17, 1149-1159, 10.5194/hess-17-1149-2013, 2013.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Verbunt, M., Gurtz, J., Jasper, K., Lang, H., Warmerdam, P., and Zappa, M.: The hydrological role of snow and glaciers in alpine river basins and their distributed modeling, *Journal of Hydrology*, 282, 36-55, 2003.

Viglione, A., Parajka, J., Rogger, M., Salinas, J. L., Laaha, G., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins; Part 3: Runoff signatures in Austria, *Hydrol. Earth Syst. Sci. Discuss.*, 10, 449-485, 10.5194/hessd-10-449-2013, 2013.

Vogel, R. M., and Fennessey, N. M.: Flow duration curves 1: New interpretation and confidence intervals, *J. Water Resour. Plan. Manage.-ASCE*, 120, 485-504, 10.1061/(asce)0733-9496(1994)120:4(485), 1994.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrological models, *Water Resources Research*, 39, 10.1029/2002WR001746, 2003.

Wagener, T., and Gupta, H. V.: Model identification for hydrological forecasting under uncertainty, *Stochastic Environmental Research and Risk Assessment*, 19, 378-387, 10.1007/s00477-005-0006-5, 2005.

Wagener, T., and Montanari, A.: Convergence of approaches toward reducing uncertainty in predictions in ungauged basins, *Water Resour. Res.*, 47, W06301, 10.1029/2010wr009469, 2011.

Weingartner, R., and Aschwanden, H.: Discharge regimes. Plate 5.2 of the *Hydrological Atlas of Switzerland*, 1989.

Westerberg, I. K., Guerrero, J. L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C. Y.: Calibration of hydrological models using flow-duration curves, *Hydrology and Earth System Sciences*, 15, 2205-2227, 10.5194/hess-15-2205-2011, 2011.

Westerberg, I. K., and McMillan, H. K.: Uncertainty in hydrological signatures, *Hydrol. Earth Syst. Sci.*, 19, 3951-3968, 10.5194/hess-19-3951-2015, 2015.

Wiesinger, T.: Schnee und Lawinen in den Schweizer Alpen. Winter 1999/2000. Wetter, Schneedecke und Lawinengefahr. Winterbericht SLF. Davos, Eidg. Institut für Schnee- und Lawinenforschung SLF, Davos, Switzerland, 89, 2006.

Winsemius, H., Schaefli, B., Montanari, A., and Savenije, H. H. G.: On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information, *Water Resources Research*, 45, W12422, 10.1029/2009WR007706, 2009.

Winsemius, H. C., Schaefli, B., Savenije, H. H. J., and Bastiaanssen, W. G. M.: Stepped calibration of a PUB model using evaporation and discharge spectral properties, *EGU 2008*, 2008,

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, *Advances in Water Resources*, 30, 1756-1774, 2007.

Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, *Water Resources Research*, 44, W09417, 10.1029/2007wr006716, 2008.

Zappa, M., and Kan, C.: Extreme heat and runoff extremes in the Swiss Alps, *Nat. Hazards Earth Syst. Sci.*, 7, 375-389, 2007.



## Tables

**Table 1: Model parameters to calibrate, a priori value ranges and additional constraints**

Symbol	Unit	Min.	Max.	Constraint	Meaning
$a_s$	mm/°C/day	1	6	$a_s < a_i$	Degree-day factor for snow
$a_i$	mm/°C/day	4	12	$a_s < a_i$	Degree-day factor for ice
$G_{\max}$	mm/d	0	2		Max. ground heat melt at snowpack basis
$Z_r$	mm	0	1500		Root zone depth
$L_{\max}$	mm/h	0	0.15		Max. deep leakage (feeding slow flow component)
$C$	-	3.3	30		Clapp-Hornberger exponent
$K_{\text{sat}}$	mm/h	0.01	500		Saturated hydraulic conductivity
$s_w$	-	0.1	0.4	$s_w < s_m$	Wilting point
$s_m$	-	0.4	0.6	$s_w < s_m$	Plant stress point
$k_{\text{slow}}$	1/h	1/(365*24)	1	$k_{\text{fast}} > k_{\text{slow}}$	Slow subsurface flux coefficient
$k_{\text{fast}}$	1/h	1/(10*24)	1	$k_{\text{fast}} > k_{\text{slow}}$	Fast subsurface flux coefficient
$k_{\text{ice}}$	1/h	1/(15*24)	1		Ice melt reservoir coefficient

**Table 2: Summary of hydrological signatures used in this paper; a reference is indicated if appropriate; for the coefficient of variation, the reference period is set to 01 May – 30 Aug. for all years. The signatures used in (Viglione et al., 2013) are highlighted; note, however, that no signature is normalized by the mean discharge as in (Viglione et al., 2013) because normalization is useful to compare different catchments but would obscure the source of mismatch between model and observation. Only signatures with positive discriminatory power  $D$  are retained for final model selection.**

Notation	Unit	Short name	Meaning	Positive $D$	Reference
$\delta_i$	$^{\circ}\text{C}(\text{m}^3/\text{s})^{-1}$	T*-Q* slope	Slope of the discharge-temperature cycle over season $i$	spring: yes	this paper
$r_Q$	-	maxmin ratio	Relative range of minimum to maximum discharge regime	no	this paper
$t_{\min}, t_{\max}$		max timing, min timing	Timing of minimum, maximum of the discharge regime	$t_{\min}$ : no $t_{\max}$ : yes	this paper
$Q_{\max}^*$	$\text{m}^3/\text{s}$	max. Q*	Maximum value of the discharge regime	no	this paper
$c_v$	-	cv melt	Coefficient of variation of melt-influenced discharge	yes	this paper
$\tau$	d	time scale	Time of autocorrelation function drop below 1/e	no	(Viglione et al., 2013)
$\rho_1$	-	AR1	Lag-1 autocorrelation of observed discharge	yes	(Winsemius et al., 2008)
$\Delta_p$	-	Pardé range	Difference between min. and max. Pardé coefficient	no	(Viglione et al., 2013)
$q_{\text{low}}, q_{\text{high}}$	$\text{m}^3/\text{s}$	low, high flow	5%, 95% percentile flow	$q_{\text{low}}$ : no $q_{\text{high}}$ : yes	(Viglione et al., 2013)
$m_{\text{FDC}}$	$\text{m}^3/\text{s}/\text{d}$	mid-FDC	Flow-duration curve slope between percentile 30, 70	no	(Viglione et al., 2013)

**Table 3: Model parameter values of the Nash-optimal parameter set for different periods and subcatchment (SC) or subcatchment-elevation band (SC+EB) set-up.**

Parameter / criterion	Unit	Nash-Opt. SC, 1998-2009	Nash-Opt. SC, 1981-1992	Nash-Opt. SC+EB, 1981-1992
Nash		0.76/0.77	0.82/ 0.76	0.84/ 0.79
NashLog		0.78/0.84	0.90/ 0.88	0.87/0.85
$B_E$		0.05/0.13	0.30/ 0.10	0.37/ 0.20
Bias		0.05/0.02	0.02/-0.04	0.00/-0.05
$a_s$	mm/°C/d	3.3	2.9	2.4
$a_i$	mm/°C/d	7.6	9.6	11.5
$G_{\max}$	mm/d	0.2	0.4	0.4
$Z_r$	mm	373	203	383
$L_{\max}$	mm/h	0.05	0.08	0.11
$c$	-	3.5	5.1	3.7
$K_{\text{sat}}$	mm/h	306	381	209
$s_w$	-	0.04	0.19	0.29
$s_m$	-	0.72	0.76	0.62
$k_{\text{slow}}^{-1}$	d	982.5	399.6	320.5
$k_{\text{fast}}^{-1}$	d	22.3	13.3	7.2
$k_{\text{ice}}^{-1}$	d	2.2	3.9	3.9

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Figure captions**

- Figure 1: Location of the Dischma catchment within Switzerland (left) and subcatchment set-up (right) with Swiss coordinates; source: (SwissTopo, 2008, 2005).
- Figure 2: Top: Mean annual temperature at mean catchment elevation computed from Davos station over periods of 5 years with a moving window and corresponding temperature lapse rate. Bottom: as top but mean annual precipitation at Davos (DAV) and Weissfluhjoch (WFJ) stations and specific discharge of Dischmabach Precipitation at WFJ is on average 184 mm/year higher than specific discharge.
- Figure 3: Top row: discharge regime (left) and air temperature regime vs discharge regime (right) highlighting the different seasons. Bottom row: discharge regime (left) and air temperature vs discharge regime for three different time periods. The regimes are obtained as the inter-annual mean of the day of the year, smoothed with a 30 day window.
- Figure 4: Discharge, air and water temperature and precipitation regime over the 3 periods 1981-1993 (P1), 1994-2003 (P2), 2004-2012 (P3) (selected periods adapted to the availability of water temperature data). The discharge and temperature regimes are obtained as the inter-annual mean of the day of the year (smoothed with a 30 day window); for precipitation, the monthly mean is shown; bottom left: air temperature regime plotted against water temperature regime for the available period.
- Figure 5: Discharge simulations for different parameter sets; the legend abbreviations stand for: SC: subcatchment set-up, SC+EB: subcatchment and elevation band set-up, P1: period 1 1983-1992, P2: 2000-2010
- Figure 6: Left: winter snow accumulation estimated from winter discharge versus from summer discharge for observed discharge and from the Nash-optimal simulation with the subcatchment set-up; center: flow-estimated snow accumulation versus simulated maximum annual snowpack; right: accumulation estimated from observed winter discharge vs snow accumulation estimated from simulated discharge or extracted from simulated snowpack.
- Figure 7: Temporal variability of selected runoff signatures studied in this paper for observed discharge (red) and three different Nash-optimal runs. For set-up abbreviations see Figure 5, for signature names Table 2.
- Figure 8: Discriminatory power  $D_i$  of the different signatures studied here. Included are also the values obtained if retaining all simulation with Nash>0.6, 0.7, 0.8 and bias<0.1.

Figure 9: Projection of the Nash performance of all parameter sets onto a single parameter axis (so-called dotty plots, Beven, 2006); the chosen parameter is the day factor for snow (in  $\text{mm}^{\circ}\text{C}^{-1}\text{d}^{-1}$ ). In grey are shown all sampled parameter sets; in red and black are shown the parameter sets that are within the limits-of-acceptability of a) the max. timing  $t_{\max}$  and  $t_{\max}$ ,  $\delta_3$  (spring T-Q slope); b) the melt coefficient of variation  $c_v$  and  $t_{\max}$ ,  $\delta_3$ ,  $c_v$ ; c) the 95% percentile flow  $q_{\text{high}}$  and  $t_{\max}$ ,  $\delta_3$ ,  $c_v$ ,  $q_{\text{high}}$ ; d) the lag-1 autocorrelation  $\rho_1$  and  $t_{\max}$ ,  $\delta_3$ ,  $c_v$ ,  $q_{\text{high}}$ ,  $\rho_1$ .

Figure 10: Simulation range spanned by all simulations that are within the limits-of-acceptability for the maximum annual timing; top row: for model set-up without elevation bands; bottom row: with elevation bands; shown is also the range spanned by all simulations with  $\text{Nash} > 0.5$  for the period 1983-1992

Figure 11: Simulation range spanned by all simulations that are, for period 1998-2009, within the limits-of-acceptability for all signatures with positive discriminatory power in Figure 8 (max. timing, cv melt, spring T-Q slope, high flow).

Figure 12: Signature values of the signatures not retained for model selection (for signature names see Table 2).

Figure 13: Comparison of observed and simulated annual discharge anomalies for the Dischma river with annual discharge anomalies from two nearby rivers with similar hydrologic regime and catchment size. The anomalies are obtained by dividing by the mean over the period considered here (1983-2009). For details on the selected rivers see (Schaeffli et al., 2013) or <http://www.hydrodaten.admin.ch> (accessed Oct. 2015).



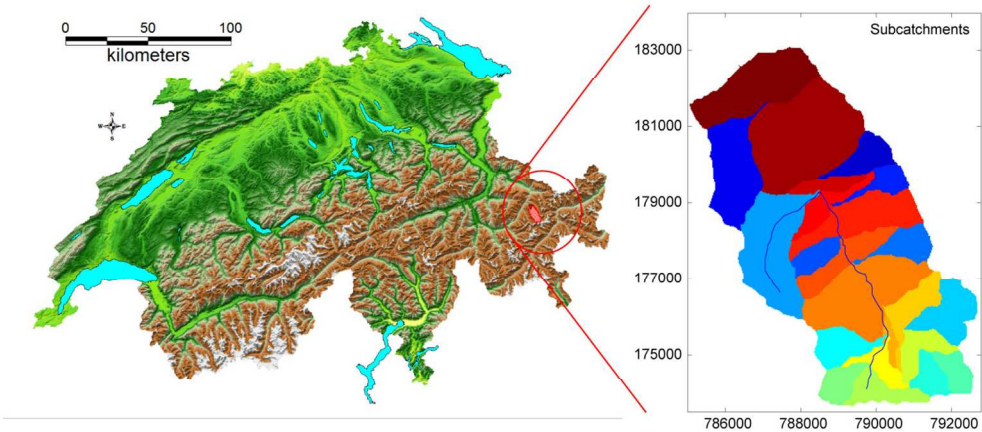


Figure 1: Location of the Dischma catchment within Switzerland (left) and subcatchment set-up (right) with Swiss coordinates; source: (SwissTopo, 2008, 2005).  
359x180mm (100 x 100 DPI)

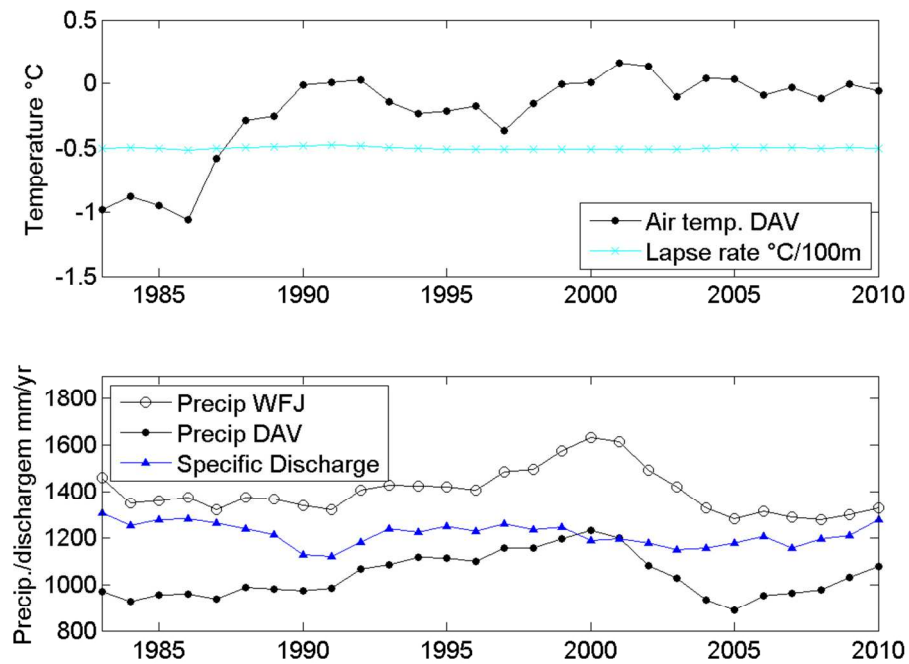


Figure 2: Top: Mean annual temperature at mean catchment elevation computed from Davos station over periods of 5 years with a moving window and corresponding temperature lapse rate. Bottom: as top but mean annual precipitation at Davos (DAV) and Weissfluhjoch (WFJ) stations and specific discharge of Dischmabach Precipitation at WFJ is on average 184 mm/year higher than specific discharge.  
203x152mm (150 x 150 DPI)

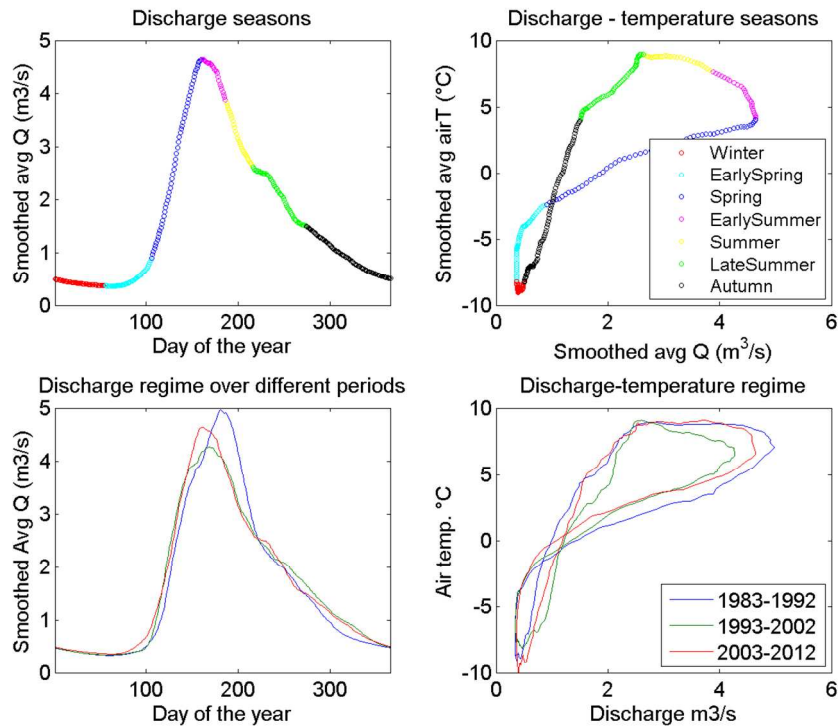


Figure 3: Top row: discharge regime (left) and air temperature regime vs discharge regime (right) highlighting the different seasons. Bottom row: discharge regime (left) and air temperature vs discharge regime for three different time periods. The regimes are obtained as the inter-annual mean of the day of the year, smoothed with a 30 day window.  
235x183mm (150 x 150 DPI)

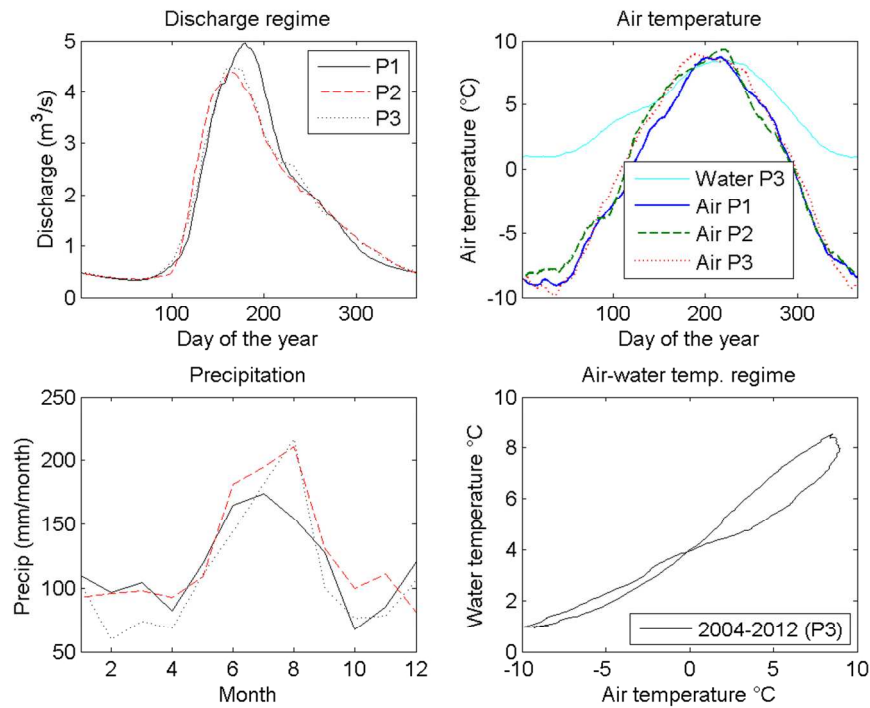


Figure 4: Discharge, air and water temperature and precipitation regime over the 3 periods 1981-1993 (P1), 1994-2003 (P2), 2004-2012 (P3) (selected periods adapted to the availability of water temperature data).

The discharge and temperature regimes are obtained as the inter-annual mean of the day of the year (smoothed with a 30 day window); for precipitation, the monthly mean is shown; bottom left: air temperature regime plotted against water temperature regime for the available period.

203x152mm (150 x 150 DPI)

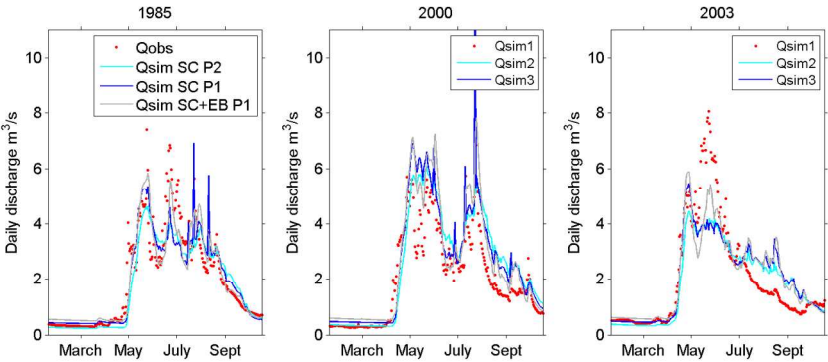


Figure 5: Discharge simulations for different parameter sets; the legend abbreviations stand for: SC: subcatchment set-up, SC+EB: subcatchment and elevation band set-up, P1: period 1 1983-1992, P2: 2000-2010  
317x119mm (150 x 150 DPI)



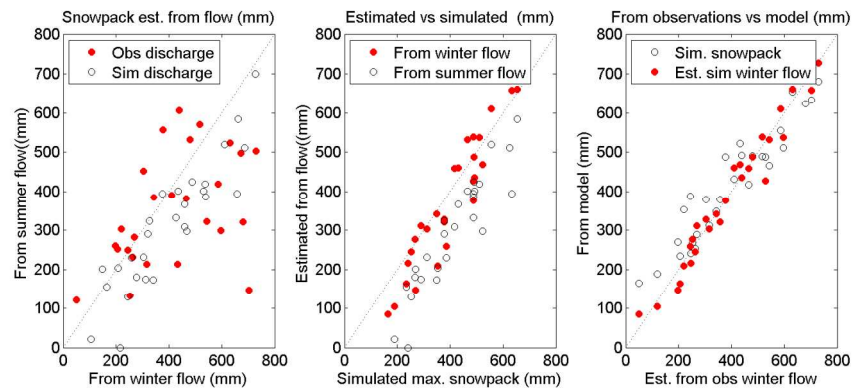


Figure 6: Left: winter snow accumulation estimated from winter discharge versus from summer discharge for observed discharge and from the Nash-optimal simulation with the subcatchment set-up; center: flow-estimated snow accumulation versus simulated maximum annual snowpack; right: accumulation estimated from observed winter discharge vs snow accumulation estimated from simulated discharge or extracted from simulated snowpack.

313x120mm (150 x 150 DPI)

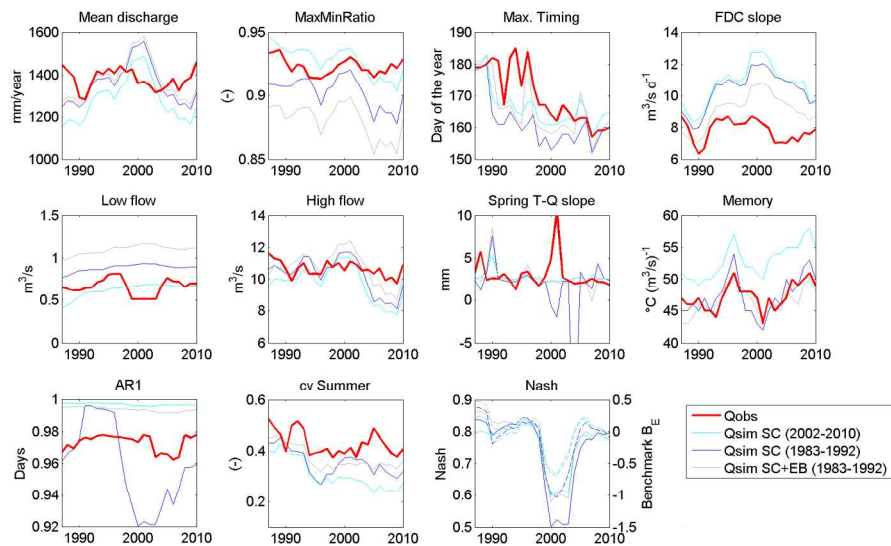


Figure 7: Temporal variability of selected runoff signatures studied in this paper for observed discharge (red) and three different Nash-optimal runs. For set-up abbreviations see Figure 5, for signature names Table 2.  
352x216mm (150 x 150 DPI)

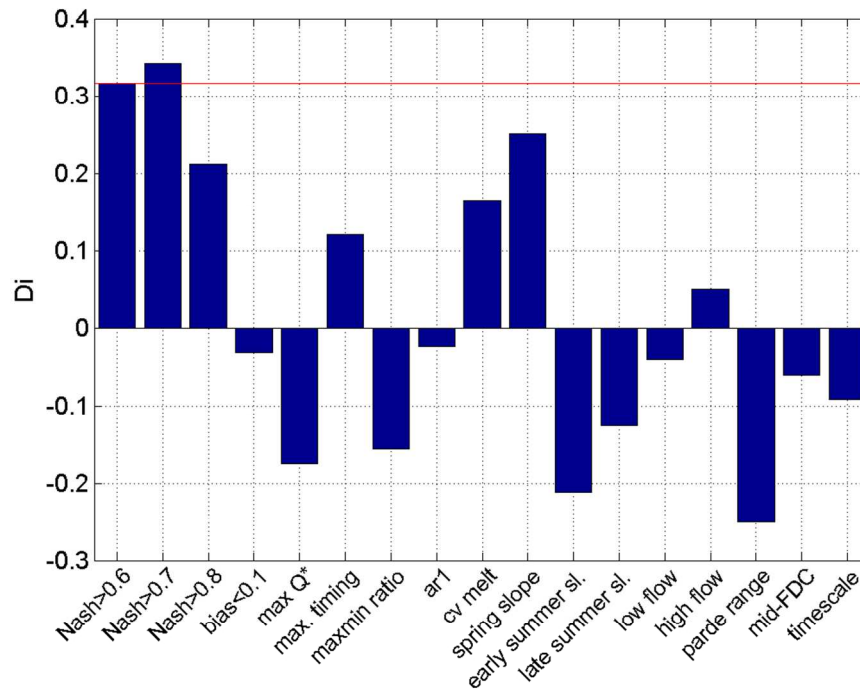


Figure 8: Discriminatory power  $D_i$  of the different signatures studied here. Included are also the values obtained if retaining all simulation with Nash>0.6, 0.7, 0.8 and bias<0.1.  
196x163mm (150 x 150 DPI)

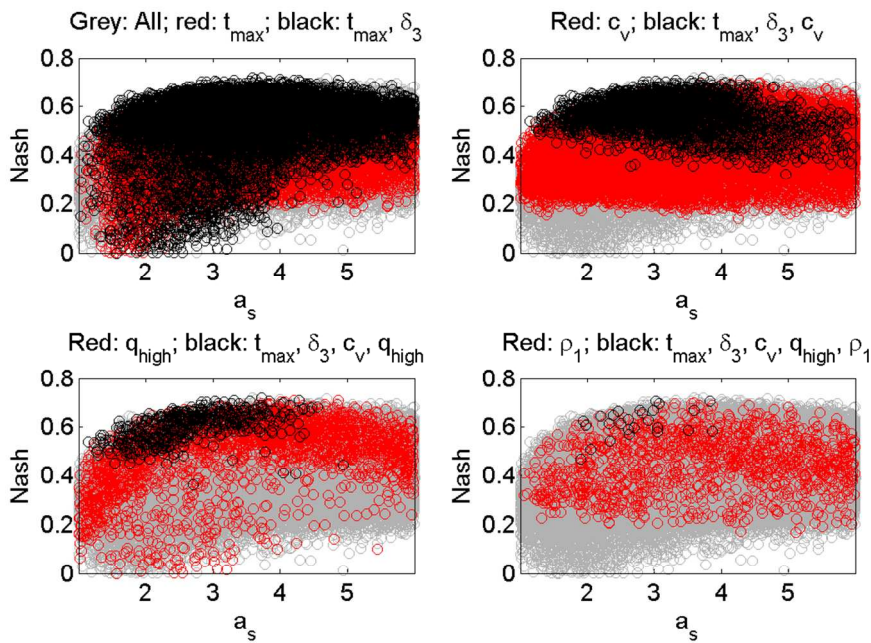


Figure 9: Projection of the Nash performance of all parameter sets onto a single parameter axis (so-called dotty plots, Beven, 2006); the chosen parameter is the day factor for snow (in  $\text{mm}^\circ\text{C-1d-1}$ ). In grey are shown all sampled parameter sets; in red and black are shown the parameter sets that are within the limits-of-acceptability of a) the max. timing  $t_{\max}$  and  $t_{\max}, \delta_3$  (spring T-Q slope); b) the melt coefficient of variation  $c_v$  and  $t_{\max}, \delta_3, c_v$ ; c) the 95% percentile flow  $q_{\text{high}}$  and  $t_{\max}, \delta_3, c_v, q_{\text{high}}$ ; d) the lag-1 autocorrelation  $\rho_1$  and  $t_{\max}, \delta_3, c_v, q_{\text{high}}, \rho_1$ .  
203x137mm (150 x 150 DPI)

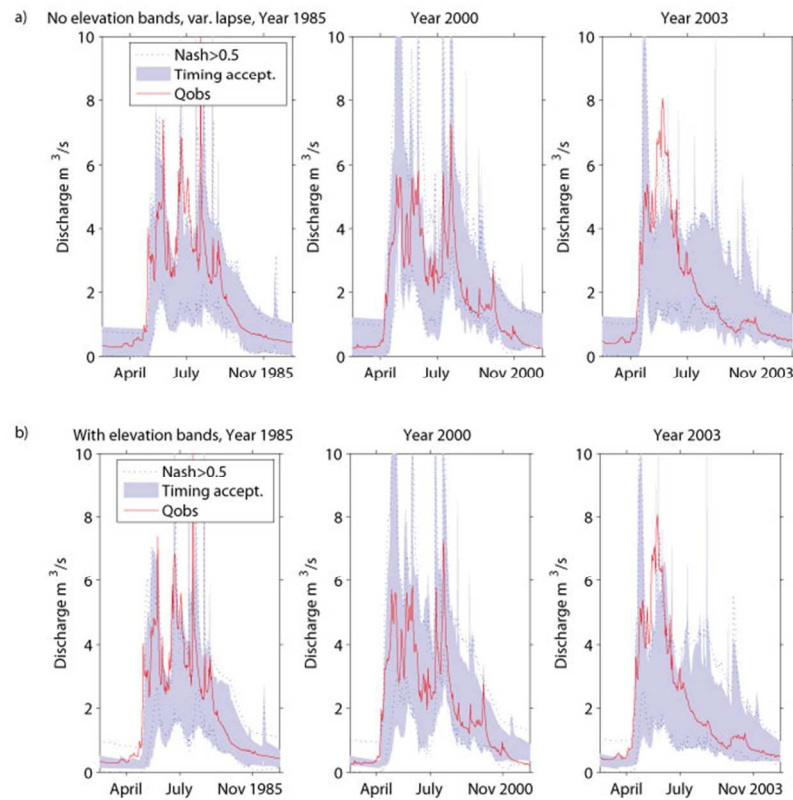


Figure 10: Simulation range spanned by all simulations that are within the limits-of-acceptability for the maximum annual timing; top row: for model set-up without elevation bands; bottom row: with elevation bands; shown is also the range spanned by all simulations with Nash>0.5 for the period 1983-1992  
215x194mm (100 x 100 DPI)



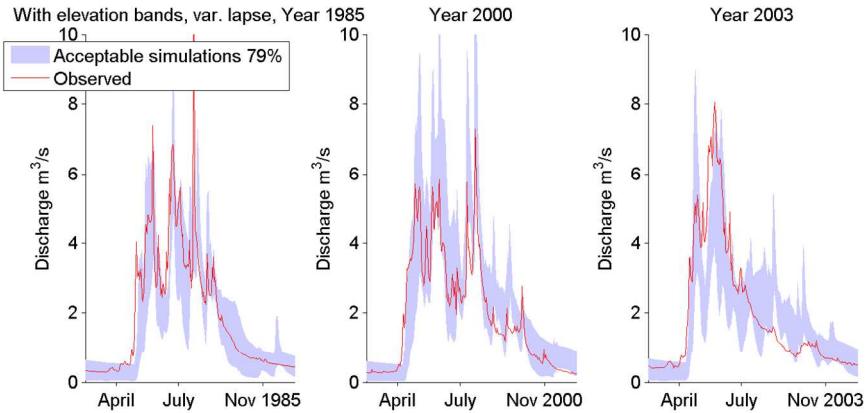


Figure 11: Simulation range spanned by all simulations that are, for period 1998-2009, within the limits-of-acceptability for all signatures with positive discriminatory power in Figure 8 (max. timing, cv melt, spring T-Q slope, high flow).  
270x119mm (150 x 150 DPI)

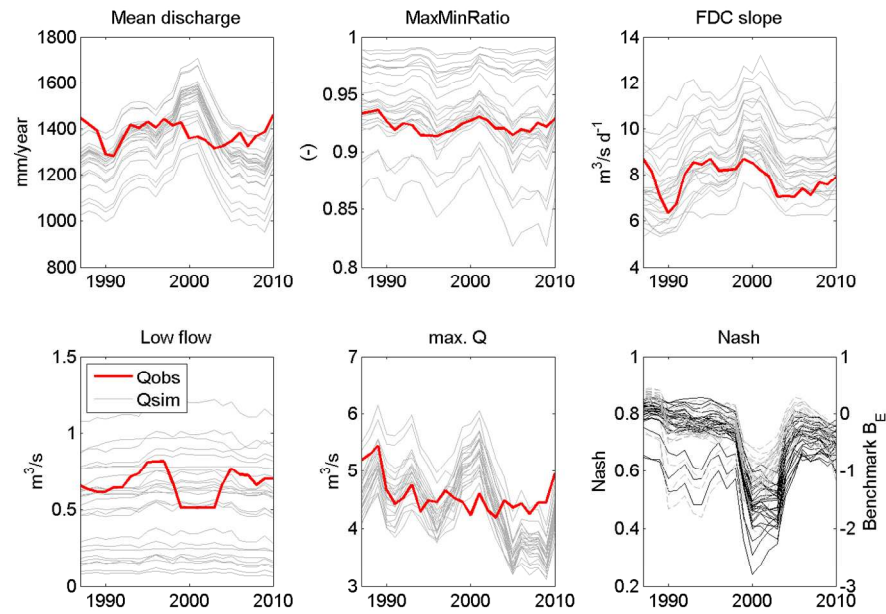


Figure 12: Signature values of the signatures not retained for model selection (for signature names see Table 2).  
270x183mm (150 x 150 DPI)

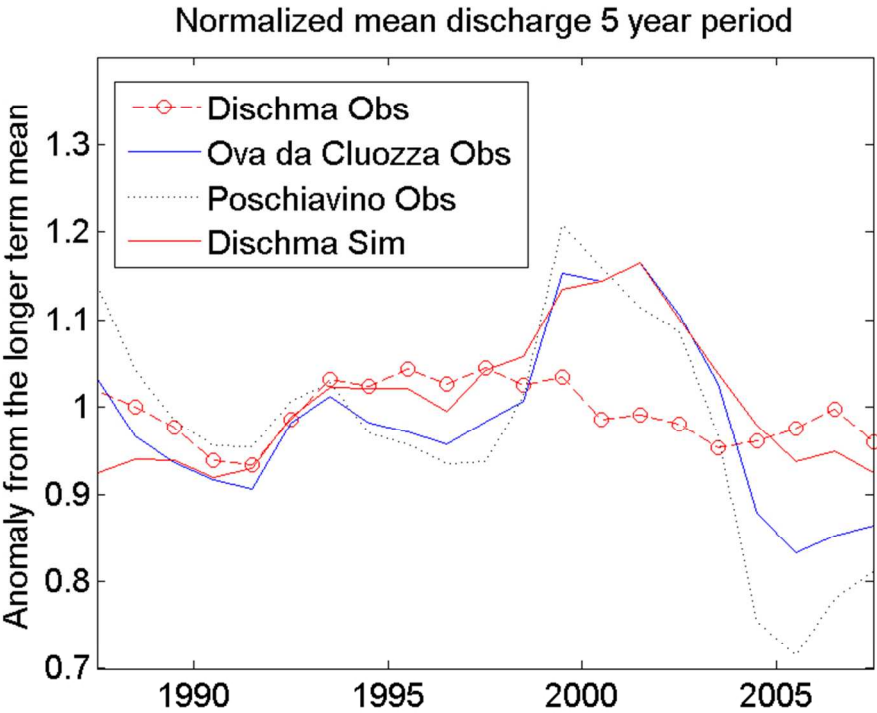


Figure 13: Comparison of observed and simulated annual discharge anomalies for the Dischma river with annual discharge anomalies from two nearby rivers with similar hydrologic regime and catchment size. The anomalies are obtained by dividing by the mean over the period considered here (1983-2009). For details on the selected rivers see (Schaepli et al., 2013) or <http://www.hydrodaten.admin.ch> (accessed Oct. 2015).  
148x111mm (150 x 150 DPI)